

## Full-HD 영상의 정수 단위 고속 움직임 예측 기법

이대현 박상욱 \*심재영 \*\*김창수 이상욱

서울대학교 \*울산과학기술대학교 \*\*고려대학교

dhlee@ipl.snu.ac.kr

## Integer-Pel Fast Motion Estimation of Full-HD sequences

Dae-Hyun Lee Sang-Uk Park \*Jae-Young Sim \*\*Chang-Su Kim Sang-Uk Lee

Seoul National University \*Ulsan National Institute of Science and Technology

\*\*Korea University

## 요약

본 논문에서는 Full-HD 영상에서 사용되는 H.264/AVC의 정수 단위 고속 움직임 예측 방법을 제안한다. 제안되는 알고리즘에서는 다중 해상도 고속 움직임 예측 기법에 기반을 두어 두 계층이 각기 탐색된다. 낮은 해상도의 계층에서는 움직임 벡터 예측자를 중심으로 좁은 탐색 영역을 2 단계로 탐색하여 최적의 점을 찾는다. 높은 해상도의 계층에서는 4 단계로 탐색을 하여 탐색점의 개수를 줄인다. 그리고 두 계층에서 각기 구해진 탐색점들의 비용을 비교하여 매크로블록의 최종 움직임 벡터를 구한다. 시뮬레이션 결과에서는 기존의 연구 결과보다 JM을 기준으로 BD-Rate는 1.55 % 높았고, BD-PSNR은 0.05 dB 낮아진 데 비해 시간은 63 % 만큼 감소하여 높은 속도를 낼 수 있었다.

## 1. 서론

H.264/AVC 비디오 압축 표준은 이전의 비디오 압축 표준들에 비해 더 나은 성능을 가지고, 방송, 멀티미디어 통신 등 각종 분야에서 핵심 기술로 자리잡고 있다[1]. H.264/AVC의 높은 성능은 다양한 블록 크기와 다수의 참조프레임, 다양한 인트라 예측 모드, 윌-오프 최적화, 새로운 엔트로피 코딩의 채택 등 여러 기술에 기반을 두고 있다. 하지만 다양한 기술의 도입으로 인해 계산량이 증가하여 실시간으로 인코딩을 하기에는 어려움이 있다. 여러 기술 중에서 가장 계산량이 많은 부분은 정수 단위 고속 움직임 예측 단계이다[2].

H.264의 정수 단위 고속 움직임 예측은 이전 표준들에서의 그것과 같아 연구가 많이 되어 왔다. 하지만 Full-HD 영상에서 정수 단위 고속 움직임 예측을 제안한 연구는 부족한 실정이다. Full-HD 영상은 차세대 방송 시스템에서 많이 확산되고 있는 1920x1080의 고품질 영상이며, 1280x720의 HD 영상보다 약 2배의 픽셀 수를 갖기 때문에 움직임 추정을 위해 시간이 더 걸린다. 또한 HD 영상의 움직임 벡터와 같은 물리적 크기의 움직임 벡터를 예측하기 위해 Full-HD 영상에서는 탐색 영역을 더 넓혀야 하기 때문에 실시간 인코딩을 위해 고속 움직임 추정을 하는 것은 필수적이다.

현재 Full-HD 영상을 기준으로 고속 움직임 예측을 한 대표적인 논문의 알고리즘인 PMRME는 다중 해상도 움직임 예측에 기반을 두어 탐색한다[3]. 가장 낮은 해상도인 계층0에서는 주변 매크로블록 움직임 벡터들의 Median으로 계산되는 움직임 벡터 예측자를 중심으로 좁은 탐색 영역을 2 단계로 탐색하여 최적의 점을 찾는다. 높은 해상도의 계층에서는 4 단계로 탐색을 하여 탐색점의 개수를 줄인다. 그리고 두 계층에서 각기 구해진 탐색점들의 비용을 비교하여 매크로블록의 최종 움직임 벡터를 구한다. 시뮬레이션 결과에서는 기존의 연구 결과보다 JM을 기준으로 BD-Rate는 1.55 % 높았고, BD-PSNR은 0.05 dB 낮아진 데 비해 시간은 63 % 만큼 감소하여 높은 속도를 낼 수 있었다.

은 영역을 탐색하여 느린 움직임을 예측한다. 중간 해상도인 계층1은 더 빠른 움직임을 예측하기 위해 좀 더 넓은 영역을 탐색하며, 계층2는 가장 넓은 영역을 탐색한다. 계층1, 2는 영점을 중심으로 하여 하드웨어로 구현되었을 때 메모리 재사용을 가능하게 한다. 하지만 계층별 탐색점의 개수가 각각 256, 1024, 4096 개로 매우 많고 계층2는 거의 선택되지 않는다.

## 2. 제안하는 알고리즘

본 논문에서 제안하는 알고리즘은 기존 연구에서의 다중 해상도 움직임 예측 방법을 이용하되 두 개의 계층만을 사용한다. 낮은 해상도인 계층0에서는 움직임 벡터 예측자를 중심으로 16x16의 좁은 영역을 탐색한다. 대부분의 움직임은 느리기 때문에 움직임 벡터 예측자 근처에 있을 확률이 높고 16x16의 영역 만으로도 대부분의 최적의 움직임 벡터를 찾을 수 있다. 그 영역에서 그림 1(a)와 같이 64개의 샘플링된 점들 중에서 예러가 최소인 점을 찾고 다시 그 주위로 8개의 탐색점을 조사하는 2 단계 탐색을 한다. 높은 해상도인 계층1은 256x256의 넓은 탐색 영역을 이용하여 빠른 움직임을 예측한다. 영역이 넓기 때문에 그림 1(b)와 같이 단계별로 16개의 탐색점을 두는 4 단계 탐색을 한다. 계층1은 알고리즘이 하드웨어로 구현되었을 때 메모리 재사용을 고려하여 영점을 기준으로 한다. 그리고 탐색점 당 계산량을 줄이기 위해 SAD를 계산하기 위한 현재블록과 참조블록을 4:1로 샘플링하고, 한 화소당 비트 깊이를 8에서 6으로 한다.

정수 단위 움직임 예측에서 각 인터모드별로 최적의 점이 계산이

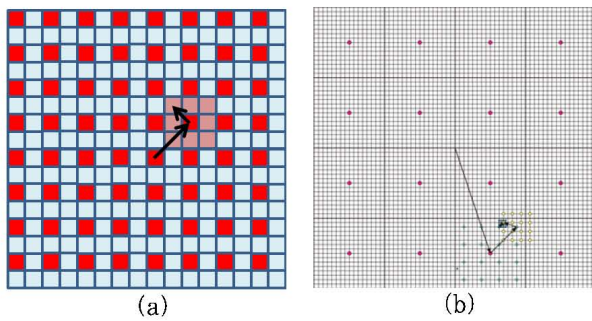


그림 1. 계층별 단계 탐색 (a) 계층0의 2단계탐색, (b) 계층1의 4단계탐색

되면, 소수 단위 움직임 예측에서는 그 인터모드들 중에서 비용이 가장 적은 두 개의 모드만을 택하여 계산된다. 그럼으로써 계산되는 블록을 기존의 41개에서 3~18 개로 줄인다.

### 3. 실험결과

실험은 Full-HD 영상인 Pedestrian Area, Crowd Run으로 하였으며, 실험에 쓰인 환경 변수는 표 1과 같다. 실험에 쓰인 측정 단위는 Bjontegaard의 것을 사용하였다[4]. 시간은 다음과 같이 정의된다.

$$Time = \frac{Time_{Algorithm}}{Time_{JM}} \times 100$$

실험한 결과, 표 2를 보면 기존 알고리즘에 비해 제안된 알고리즘이 JM기준으로 BD-Rate는 1.55 % 높고, BD-PSNR은 0.05 dB 낮아졌지만 시간은 1.19 만큼, 즉 제안된 알고리즘이 기존 알고리즘보다 63 % 만큼 감소하여 율-왜곡 성능 대비 높은 시간 감소를 나타내었다. 그림 2에서 보여주는 율-왜곡 그래프에서도 두 알고리즘 사이의 성능 하락은 거의 눈에 띄지 않았다.

### 4. 결론

본 논문에서는 Full-HD 영상에서의 정수 단위 고속 움직임 예측을 위한 방법을 제안하였다. 두 개의 계층을 갖는 다중 해상도 움직임을 위한 방법을 제안하였다.

표 1. 실험 환경 변수

환경 변수	값
Source Size	1920x1080
QP	20, 24, 28, 32
Search Range	128
Frame Structure	IPPP
Number of Frames	100
Frame Rate	50
Number of Reference Frames	1

표 2. 기존 연구와 제안된 알고리즘 간의 비교

Sequence	Algorithm	BD-Rate(%)	BD-PSNR(dB)	Time
Pedestrian Area	PMRME	1.01	-0.03	2.06
	Proposed	3.38	-0.09	0.87
Crowd Run	PMRME	0.65	-0.04	1.81
	Propose	1.38	-0.07	0.61
Average	PMRME	0.83	-0.03	1.93
	Proposed	2.38	-0.08	0.74

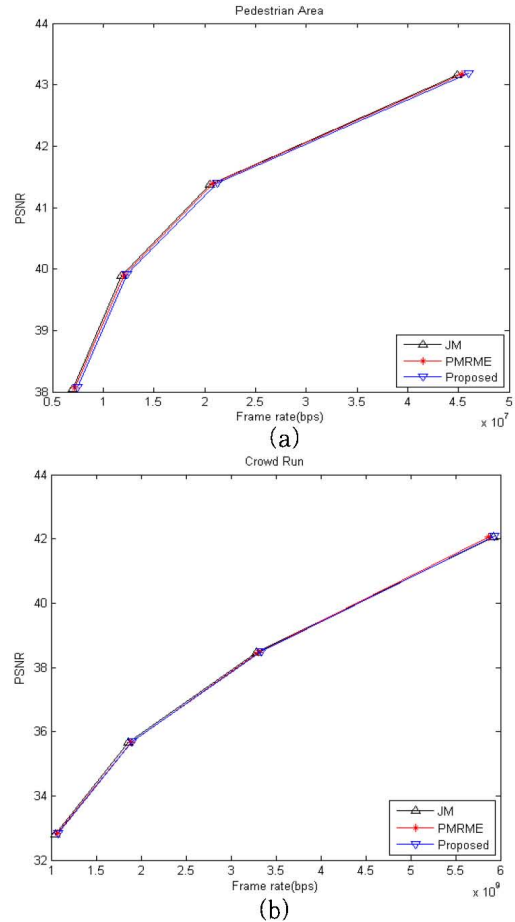


그림 2. Full HD 영상의 율-왜곡 곡선 (a) Pedestrian Area, (b) Crowd Run

예측 기법에 기반을 둔다. 낮은 해상도의 계층은 좁은 탐색 영역에서 움직임 벡터의 대다수를 차지하고 있는 느린 움직임을 예측하고, 높은 해상도의 계층은 넓은 탐색 영역에서 소수의 빠른 움직임 벡터를 예측한다. 실험 결과 기존의 연구 결과보다 JM을 기준으로 BD-Rate는 1.55 % 높았고, BD-PSNR은 0.05 dB 낮아진데 비해 시간은 만큼 다시 말해 63 % 만큼 감소하여 높은 속도를 낼 수 있었다.

### 참고 문헌

[1] ISO/IEC ITU-T Rec. H.264: Advanced Video Coding for Generic Audiovisual Services, Joint Video Team(JVT) of ISO-IEC MPEG & ITU-T VCEG, Int. Standard, May 2003.  
 [2] T.-C. Chen, S.-Y. Chien, Y.-W. Huang, C.-H. Tsai, C.-Y. Chen, T.-W. Chen, and L.-G. Chen, "Analysis and Architecture Design of an HDTV720p 30 Frames/s H.264/AVC Encoder," International Trans. CSVT, vol. 16, no. 6, pp. 673-688, June 1989.  
 [3] C.-C. Lin, Y.-K. Lin, and T.-S. Chang, "PMRME: A Parallel Multi-Resolution Motion Estimation Algorithm and Architecture for HDTV sized H.264 Video Coding," ICASSP, April 2007.  
 [4] G. Bjontegaard, "Calculation of Average PSNR differences between RD curves," in Joint Video Team(JVT) of ISO/IEC MPEG and ITU-T VCEG, Mar. 2001, Doc. VCEG-M33.